PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, C12N 15/31, C07K 14/21, 16/12 — A3

(11) International Publication Number: WO 97/01647

(43) International Publication Date: 16 January 1997 (16.01.97)

(21) International Application Number: PCT/US96/11132

(22) International Filing Date: 28 June 1996 (28.06.96)

(30) Priority Data:
60/000,828 28 June 1995 (28.06.95) US

(71) Applicant: TRUSTEES OF HEALTH & HOSPITALS OF THE CITY OF BOSTON, INC. [US/US]; 6th floor, 1010 Massachusetts Avenue, Boston, MA 02119 (US).

(72) Inventor: GOLDSTEIN, Richard; 12 Howland Street, Cambridge, MA 02138 (US).

(74) Agents: PIERRI, Margaret, A. et al.; Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
01 May 1997 (01.05.97)

(54) Title: DNA SEQUENCES FOR IDENTIFYING HIGHLY TRANSMISSIBLE LINEAGES OF PSEUDOMONAS (BURKHOLDERIA) CEPACIA

(57) Abstract

The present invention provides DNA based fingerprints and DNA sequences for indentifying highly transmissible lineages of *Pseudomonas cepacia*. More specifically, the present invention provides genetic band patterns or DNA based fingerprints for identifying highly transmissible lineages of *Pseudomonas cepacia* produced by: (a) ribotyping analyses, i.e., the determination of restriction fragment length polymorphisms (RFLPs) associated with the multicopy RNA operon (rrn); and (b) pulsed field gel electrophoresis (PFGE)-based resolution of chromosomal macro-RFLP patterns. Also provided are unique primer oligonucleotide sequences and DNA probes derived from variants of a gene, encoding a 17-KDa major subunit pilin protein (cblA) of a cystic fibrosis-associated *Pseudomonas cepacia*. The invention also discloses methods and diagnostic kits for identifying highly transmissible lineages of *Pseudomonas cepacia* that utilize the above DNA based fingerprints, oligonucleotide primers and DNA probes. Another object of this invention is to provide isolated DNA molecules encoding a 17-KDa major subunit pilin protein of certain cystic fibrosis-associated *Pseudomonas cepacia* strains, as well as recombinant DNA molecules, transformed hosts and methods for the production of that protein. Also contemplated are antibodies to the 17-KDa major subunit pilin protein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

DNA SEQUENCES FOR IDENTIFYING HIGHLY TRANSMISSIBLE LINEAGES OF PSEUDOMONAS (BURKHOLDERIA) CEPACIA

FIELD OF THE INVENTION

This invention relates generally to the detection of highly transmissible strains of *Pseudomonas (Burkholderia) cepacia*,* and particularly to the use of DNA based fingerprints and sequences for identifying such epidemic lineages.

BACKGROUND OF THE INVENTION

Pseudomonas cepacia is an aerobic, gram-negative bacillus with an ubiquitous distribution in soil and water. In recent years this organism has emerged as an important pathogen among cystic fibrosis (CF) patients. CF patients with respiratory colonization or infection with *P. cepacia* have higher morbidity and mortality than those CF patients not infected by this organism.[1-6] While the significant increase in *P. cepacia* infection suggests epidemic spread[3, 6-8], the source and transmissibility of

---

* In view of recent molecular phylogenic analyses, the bacterium *Pseudomonas cepacia* has been renamed *Burkholderia cepacia*. However, in order to be consistent with the terminology of the literature cited in the specification, the original nomenclature is maintained herein.

P. cepacia remains controversial[9]. Nonetheless, given the potentially grave consequence of *P. cepacia* infection, stringent infection control policies have been adopted, many CF camps in North America have been closed and all but one lung transplant center have ceased to accept *P. cepacia*-infected CF patients as transplant candidates.

The epidemiology of *P. cepacia* infection has been previously examined by both ribotyping[1, 9] and pulsed-field gel electrophoresis (**PFGE**)-based resolution of chromosomal macro-restriction fragment profiles[3, 9, 10]. Comprehensive studies applying both methods generated two very different conclusions regarding clonality, persistence, and transmissibility. One study in the UK (Western General Hospital, Edinburgh), found a clonal relationship among isolates from 13 patients over six years[3]. In contrast, during an eight year period at a US CF center (UNC Hospitals, Chapel Hill), not a single identical or closely-related strain was found among 23 infected clinic and lung transplant patients[9]. Serial isolate analysis further confirmed this picture, typically demonstrating persistent infection by a single strain.

There also existed an isolate collection from another CF center (Hospital for Sick Children, Toronto) where there was anecdotal evidence for an epidemic of *P. cepacia*. Although the isolates were not characterized for genetic-relatedness, they had been uniformly resolved to express peritricious, giant cable (**Cbl**)-like pili that specifically bound to CF mucin and airway epithelial cells[4, 11].

As the *cblA* pilin subunit gene encoded by all 15 of the Toronto isolates was the first adhesin pilus

subunit gene characterized for *P. cepacia*, a subsequent hybridization-based survey for the presence of *cblA* was carried out on multiple isolates from seven other CF centers in the US (Chapel Hill NC, Jackson MS, Norfolk VA, Cleveland OH, Phila. PA, New York NY) and Europe (Edinburgh) as well as clinical and environmental strains. All of these isolates were $cblA^-$ except for one isolate from a CF center in Jackson Mississippi.

By examining the genetic-relatedness of $cblA^+$ and $cblA^-$ strains, we have discovered the emergence of a highly transmissible lineage, seemingly adapted for efficient transmission in the CF population. Resolved genetic markers uniquely associated with this lineage, which may be used to rapidly identify its presence, are of immediate practical importance in CF centers in both Europe and North America.

SUMMARY OF THE INVENTION

It is an object of this invention to provide chromosomal restriction fragment length polymorphism (RFLP) patterns (generically referred to as DNA based fingerprints) and DNA sequences for identifying highly transmissible lineages of *Pseudomonas cepacia*.

More specifically, the present invention provides DNA based fingerprints for identifying highly transmissible lineages of *Pseudomonas cepacia* produced by: (a) ribotyping analyses, i.e., the determination of RFLPs associated with the multicopy RNA operon(rrn); and (b) pulsed field gel electrophoresis (PFGE)-based resolution of chromosomal macro-RFLP patterns.

Another object of this invention is to provide isolated DNA molecules encoding a 17-KDa major

subunit pilin protein (cblA) of cystic fibrosis-associated, epidemic and non-epidemic *Pseudomonas cepacia* strains. Also contemplated is the recombinant production of polypeptides using the isolated DNA molecules. The recombinant polypeptides are then used to produce antibodies for use in methods for identifying epidemic strains of *Pseudomonas cepacia*.

It is also an object of this invention to provide unique primer oligonucleotide sequences derived from *cblA* gene variants for use in polymerase chain reaction (PCR)-based methods for detection of highly transmissible strains of *Pseudomonas cepacia*.

It is also an object of this invention to provide DNA probes derived from unique regions of variant *cblA* gene sequences for detection of highly transmissible strains of *Pseudomonas cepacia*.

It is also an object of this invention to provide methods for identifying highly transmissible lineages of *Pseudomonas cepacia* that utilize the above DNA based fingerprints, oligonucleotide primers and DNA probes.

It is a further object of this invention to provide diagnostic kits for identifying highly transmissible strains of *Pseudomonas cepacia*.

Brief Description of the Drawings

Figs 1 A-C together depict RFLPs of 17 *P. cepacia* isolates cited in the specification. Lane order for the 17 isolates is maintained in all three figures. Isolate numbers of examined strains appear at the top of the figure immediately above each lane. Subscript letters preceding isolate number indicate CF center from which *P. cepacia* (PC) strain was isolated: PC_E, Edinburgh Scotland; PC_{MS}, Jackson MS; PC_{NC}, Chapel Hill NC; PC_{NY}, New York NY; PC_{OH}, Cleveland OH; PC_T, Toronto Canada, and PC_{VA}, Norfolk VA.

Fig. 1A.: PFGE-resolved *Spe*I RFLPs. As described previously[9], samples were prepared and restriction fragments separated by pulsed field gel electrophoresis with a CHEF Mapper system (BIO-RAD) through 1% agarose using a field strength of 6 V/cm and an initial and final pulse time of 1.2 sec and 54 sec, respectively. Fragment sizes were determined using a λ concatenate ladder. Bar-code format translation of chromosomal fingerprint profiles was made using a Macintosh Quadra 950 running Gene Construction Kit (Texto). Fragments below 100 kbp are not shown. In the latter range, Toronto and Edinburgh isolates displayed in **lanes 1-8** had two identical fragments (60 kbp and 48 kbp). Other isolates (**lanes 9-17**) had polymorphic sets of three to six fragments in this lower range.

Fig. 1B: *rrn Eco*RI RFLPs. Southern blot hybridization methods were as we described previously[9, 26] using a ^{32}P-labeled *rrnB* probe spanning the entire *rrnB* operon of *E. coli* K12.

Fig. 1C: *cblA* hybridization analysis of *Eco*RI generated RFLPs. This was accomplished by stripping bound *rrn* probe from the membrane used in **Fig. 1B**, followed by hybridization with a previously described *cblA* gene probe[4] using standard methods[26, 27].

Fig. 2: *rrn*-RFLP based phylogenic tree of representative isolates from patients at seven CF centers in North America (Chapel Hill NC, Jackson MS, Norfolk VA, Cleveland OH, Philadelphia PA, New York NY, Toronto Ontario) and Europe (Edinburgh) plus environmental and clinical (non-CF) sources. All cited isolates are described in the text and in the Methods section. Indicated isolate number is followed by source (CF, environmental or clinical). $cblA_1^+$, $cblA_2^+$, isolate(s) that encode the *cblA* gene (**Fig. 1C**) and express adhesin Cbl pili (**Fig. 3**). $cblA_1^+$, identical 501 bp sequence carried by Toronto and Edinburgh CF center isolates (**Fig. 4**); $cblA_2^+$, polymorphic 501 bp sequence carried by Jackson Mississippi CF center isolate PC_{MS}_2323 (**Fig. 4**). Number above each branch indicates the percentage of time each was joined together under bootstrap analysis[14] (confidence intervals less than 10 have been omitted for clarity). The lineages included in this tree are representative of the larger sample of isolates collected. Multiple CF patient serial isolates of an identical *rrn* RFLP profile have not been included as they do not affect the tree topology. However, multiple isolates from Toronto (PC_T_5, PC_T_7) and Edinburgh (PC_E_1359, PC_E_2315) CF centers are noted because further analysis by DNA sequence revealed that

the *cblA* genes encoded by these four isolates are identical (**Fig. 4**). The remaining 13 and 11 isolates, respectively, from each of these two CF centers are members of the indicated epidemic lineage based on 100% correlation of their *rrn* RFLP profiles with those of the prototypic patterns of the Toronto/Edinburgh isolates shown in **Fig. 1B**.

Fig. 3: Transmission electron micrograph of Toronto epidemic strain PC_{T}-7 expressing **Cbl** adhesin pili. High resolution was achieved with a JOEL 100 CX electron microscope as previously described[11]. Bar in lower right corner, 0.1 μm.

Fig. 4: Identical 501 bp sequence (top-most line) of the *cblA* structural gene encoded by 2 prototypic Toronto epidemic isolates (PC_{T}-7 and PC_{T}-5) and 2 prototypic Edinburgh isolates (PC_{E}-2315 and PC_{E}-1359) compared to the variant *cblA* sequence carried by the single Jackson Mississippi CF center isolate PC_{MS}-2323 (lower line).

DETAILED DESCRIPTION OF THE INVENTION

One aspect of the present invention is directed to the identification of highly transmissible strains of *Pseudomonas cepacia* using DNA based fingerprints. More specifically, such identification

is accomplished by the use of fingerprints produced by: (a) ribotyping analyses, i.e., the determination of restriction fragment length polymorphisms (RFLPs) associated with the multicopy RNA operon (rrn); and (b) pulsed field gel electrophoresis (PFGE)-based resolution of chromosomal macro-RFLP patterns.

The techniques for ribotyping are known to those of skill in the art. Generally, DNA samples from bacterial strains are isolated, digested with a restriction endonuclease such as *Eco*R1 and separated by agarose gel electrophoresis. The DNA fragments from the agarose gel are then transferred to nitrocellulose membranes and probed with either radiolabeled or chemiluminescent *E. coli* ribosomal RNA (*rRNA*) probes. The RFLPs detected by hybridization with the *E. coli rRNA* probes are then analysed to categorize the bacterial strains according to their distinctive bands of *rRNA* encoding DNA (i.e., "DNA fingerprints").

Similarly, the skilled artisan would also be familiar with the published methods for PFGE-based resolution of chromosomal macro-RFLP patterns. Such methods involve immobilization of bacteria in agarose plugs after which the bacteria are lysed and the DNA is digested with an infrequently (rare) cutting restriction endonuclease such as *Spe*I. The agarose plugs with the digested DNA are then subjected to transverse alternating field electrophoresis in agarose gels using commercially available equipment. The gels are generally stained with ethidium bromide or cyber-green and the resulting chromosomal RFLPs are analysed preferably after computer-generated translation of the RFLP profiles into bar code format using commercially available equipment and software.

Also well known to those skilled in the art are the standard criteria for the statistical analysis of the ribotypes and the chromosomal RFLP profiles obtained using the above methods. Using these statistical criteria, bacterial strains may be compared and their genetic relationships characterized. In the present invention, the determination of whether a given strain of *Pseudomonas cepacia* is highly transmissible is made by comparing its ribotypic and chromosomal RFLP profiles to those of a known strain of highly transmissible lineage.

In another aspect of the present invention, isolated DNA molecules encoding a 17-KDa major subunit pilin protein (cblA) of cystic fibrosis-associated, epidemic and non-epidemic *Pseudomonas cepacia* strains are provided. The genes encoding the cblA protein are isolated and sequenced by standard techniques. The *cblA* genes isolated from epidemic and non-epidemic strains of *Pseudomonas cepacia* enabled for the first time the comparison of those variant genes and the identification of the differences in their nucleotide sequences.

The isolated DNA molecules are then used for the recombinant production of the 17-kDa major subunit pilin protein (cblA) of cystic fibrosis-associated, epidemic and non-epidemic *Pseudomonas cepacia* strains. As such, the present invention also contemplates recombinant DNA molecules containing the above DNA molecules and unicellular hosts transformed with those recombinant DNA molecules. The recombinant polypeptides and their fragments are then used for the production of antibodies that can distinguish epidemic and non-epidemic strains of *Pseudomonas cepacia* in

standard immunologic assays such as ELISA, radioimmunoassay and western blots. The methods for recombinant protein production, protein purification and generation of antibodies are well within the purview of the ordinary skilled artisan.

This invention also provides unique primer oligonucleotide sequences derived from *cblA* gene variants for use in polymerase chain reaction (PCR)-based methods for detection of highly transmissible strains of *Pseudomonas cepacia*. These unique primers are usually synthesized using standard procedures following identification of desirable nucleotide sequences based on the comparison of the *cblA* gene sequences of epidemic and non-epidemic strains of *Pseudomonas cepacia*. The PCR techniques contemplated by and used in the present invention are well known. Essentially, the primers, DNA from the bacterial strain to be tested and a thermostable PCR enzyme are mixed and the reaction carried out according to established procedure in a thermocycler block. The PCR products are then analyzed generally by electrophoretic separation.

The invention also provides DNA probes derived from unique regions of variant *cblA* gene sequences that may be used in standard hybridization based assays such as colony hybridization or Southern blot transfers to detect highly transmissible strains of *Pseudomonas cepacia*.

The present invention also contemplates the use of the above DNA based fingerprints, oligonucleotide primers, DNA hybridization probes, polypeptides and antibodies in diagnostic kits for the

detection of highly transmissible lineages of *Pseudomonas cepacia*.

In order that this invention may be better understood, the following examples are set forth. These examples are for purposes of illustration only, and are not to be construed as limiting the scope of the invention in any manner.

EXAMPLES

Materials and Methods

Bacterial isolates:

133 *Pseudomonas cepacia* isolates were obtained from the following sources: (i) sixty-five isolates from patients at the University of North Carolina Cystic Fibrosis Center (1985 through 1993) including 17 clinic and 5 transplant patients, 4 of whom were infected transfers from other distant locations as described previously[9, 11]. Those cited in the Figures include isolates PC_{NC}-566, PC_{NC}-1711, PC_{NC}-1845, PC_{NC}-1910, PC_{NC}-1948, PC_{NC}-2008, PC_{NC}-2028, PC_{NC}-2211, PC_{NC}-2225 and PC_{NC}-2747, each from different local clinic patients, and four isolates from infected transfer patients (PC_{NY}-792 from New York, New York, PC_{VA}-1963 from Norfolk Virginia, PC_{OH}-2034 from Cleveland Ohio and the *cblA*$^{+}$ isolate PC_{MS}-2323 from Jackson Mississippi); (ii) 8 ATCC clinical and environmental isolates: human endocarditis PC_{ATCC}-13945, human urinary tract PC_{ATCC}-17765, human bronchial PC_{ATCC}-25609, human tibia fracture PC_{ATCC}-27515, forest soil (Trinidad) PC_{ATCC}-17759, onion PC_{ATCC}-25416 (the ATCC *P. cepacia* type strain), soil (California) PC_{ATCC}-29352, cornfield soil (New Jersey)

PC_{ATCC}-39277, all as described previously[9, 11]; (iii) 15 *cblA*$^+$ isolates from 15 CF center clinic patients at the Hospital for Sick Children (1987 through 1988), Toronto. Those cited in the Figures include isolates PC_T-5, PC_T-7, PC_T-19, PC_T-25[4, 11]; (iv) two isolates, PC_E-SBC27 and PC_E-SBC29, from two CF patients at Western General Hospital (Edinburgh Scotland) that were not associated with epidemic transmission within this CF center[3, 4, 11]; (v) four isolates from clinic patients at Western General Hospital (1989 through 1990), Edinburgh, all which had been associated with epidemic transmission within this CF center[3]. Those cited in the Figures include isolates PC_E-509, PC_E-1359, PC_E-1392 and PC_E-2315; (vi) ten isolates from 10 patients at a Philadelphia Pennsylvania CF center[2]. Those cited in the Figures include isolates PC_{PA}-535, PC_{PA}-542 and PC_{PA}-544; (vii) twenty-four isolates from 24 patients at the Rainbow Babies and Children's Hospital, Cleveland Ohio[17] including cited isolate PC_{OH}-524, PC_{OH}-525 and PC_{OH}-2034; and(viii) 5 bronchial isolates from 5 ventilator patients at Boston City Hospital (Massachusetts) including cited isolate PC_{MA}-3137.

Restriction Fragment Length Polymorphism Analyses

Ribotyping:

Chromosomal DNA was prepared using the following procedure. Overnight cultures were diluted 10-fold in 10 ml of LB and incubated at 37°C until they reached mid-log phase. Cells were pelleted, washed twice with 0.9% NaCL, and then resuspended in cold TE (10 mM Tris-HC1, 10 mM EDTA [pH 8.0]). Lysozyme was added to a final concentrations of 20 mg/ml, and the

solution was incubated at 37°C for 30 min. Proteinase K (20 mg/ml in TE; Sigma) and sodium dodecyl sulfate were added to final concentration of 20 µg/ml and 0.1%, respectively. The lysates were incubated overnight at 50°C. Sarcosine-free acid (Sigma) was added to a final concentration of 2%, and the solution was mixed gently. The DNA was then purified by cesium chloride-ethidium bromide equilibrium density gradient centrifugation.[20]

*Eco*RI rRNA (*rrn*) restriction fragment length polymorphisms (RFLPs) for phylogenetic analyses were analyzed by using an *rrnB* probe spanning the entire *rrnB* operon of *Escherichia coli* K-12.[28] Plasmid DNA was purified by the alkaline lysis method[29], followed by cesium chloride-ethidium bromide gradient centrifugation.[20] The DNA restriction fragment used to generate the probe was separated by horizontal slab gel electrophoresis in 0.8% low-melting point agarose. The relevant restriction fragment from a slice of the gel was radiolabeled in the agarose by random oligonucleotide priming with [α-^{32}P]dCTP (Dupont, NEN Research Products).[30]

The chromosomal DNA prepared according to the above-described method was restricted with *Eco*R1 and the fragments were separated by agarose gel electrophoresis. Following electrophoresis, restriction fragments were transferred to nitrocellulose membranes and hybridized with the *rrn*B probe described above. Hybridized bands were visualized by autoradiography and the banding patterns were analysed using standard methods.[9,26]

Pulsed-field gel electrophoresis (PFGE):

Cells were grown to early log phase in LB, harvested by centrifugation, washed with 1:1 TE buffer, resuspended in 1 ml of 10:1 TE buffer, and mixed with 1.2 volumes of melted 1% InCert agarose (Bio-Rad) in TE buffer. The mixture was dispensed into 120-µl insert molds (Pharmacia) and allowed to solidify on ice. Plugs were sliced and incubated in 20 µg of lysozyme per ml at 37°C for 12 h. The lysozyme buffer was replaced with ESP buffer (0.5 MEDTA [pH9], 1% sarcosyl, 200 µg of proteinase K [Sigma] per ml), and the plugs were incubated at 37°C for 5 h and then washed with TE buffer for 4 h at 37°C. Single plug slices were incubated with *SpeI* (Boehringer Mannheim) in restriction enzyme buffer for 2 h. Restriction fragments were separated by PFGE using a CHEF Mapper system (Bio-Rad) through 1% agarose (Bio-Rad; Molecular Biology Certified grade) in 0.5 X TBE (1 X TBE is 0.9 M Tris-HCl, 0.9 M boric acid, and 1.0 mM EDTA). Electrophoresis was performed for 24h in 0.5 X TBE buffer at 14°C, with a field strength of 200 V (6V/cm) and initial and final pulse times of 1.2 and 54 s, respectively. Lambda concatemers were used as DNA size standards. The gels were stained with ethidium bromide or cyber-green and photographed under UV transillumination with a Polaroid camera. Fragment sizes were determined, and the computer-generated translation of chromosomal fingerprint profiles was made into bar code format by using a Macintosh Quadra 950 and Gene Construction Kit (Texto) software.

Statistical Analysis of Ribotypes and PFGE Patterns:

According to established criteria for *P. cepacia*[24], strains were assigned to the same ribotype

when comparison of sizes of hybridizing fragments revealed 3 or fewer bands differing between the two patterns under comparison.

Standard criteria were used for comparing PFGE chromosomal fingerprints.[23] Briefly, PFGE patterns were considered different (i) when they had the same number of DNA fragments but when the size of at least one band varied by more than I standard deviation (5%), (ii) when they exhibited a different number of DNA fragments, or (iii) when the sum of the sizes of the differing bands in the first PFGE pattern did not correspond to that of the differing DNA fragments in the second PFGE pattern.

Quantitative pairwise comparison of both types of RFLP patterns was accomplished using the Dice coefficient of similarity calculated as $D = 2n_{xy}/(n_1 + n_2)$, where n_1 is the total number of DNA fragments from strain X, n_2 is the total number from strain Y, and n_{xy} the number of fragments identical in the two strains[12, 25]. The coefficient of similarity for two PFGE RFLPs $D \geq 0.90$ represents closely related strains, while unrelated strains have $D \leq 0.60$. Intervening values, remarkably, were not observed, and values between 0.5 and 0.6 are rare[9]. For *rrn* RFLPs, given that *P. cepacia* strains typically display 7-10 distinct hybridizing bands, the shared ribotype (above) would correspond to D=0.79 to 0.85. Comparisons between mean values were performed by Student's t test using a Systat program (Systat Inc.).

Amplification and Sequencing of the *cblA* Gene:

Using previously described methods[9], CsCl equilibrium density gradient purified chromosomal DNA[20] was isolated from the two prototypic Toronto epidemic

isolates (PC_T_7 and PC_T_5), the two prototypic Edinburgh epidemic isolates (PC_E_2315 and PC_E_1359) and the single Jackson Mississippi CF center isolate PC_{MS}_2323. From each of these *cblA* probe-positive chromosomes, the *cblA* gene was PCR amplified[21] using a DNA thermocycler (Perkin-Elmer) with a GeneAmp PCR Core Reagents Kit (*ibid.*). Based on the previously determined sequence of the *cblA* gene encoded by isolate PC_T_74, sense and antisense primers used for these reactions were 5'-CCAAAGGACTAACCCA-3' and 5'ACGCGATGTCCATCACA-3', respectively. PCR reactions were: ~~cycle 1~~, 2 min at 94°C, 2 min at 37°C, 1 min at 72°C. The remaining 29 cycles were: 1 min at 94°C, 1 min at 45°C, 1 min at 72°C, followed by 7 min extension at 72°C. PCR products were electrophoresed through 0.8% agarose and for each a single band was observed with ethidium bromide staining. Bands were electroeluted into DEAE membrane (Schleicher & Schuell) and cloned with a TA Cloning Kit (Invitrogen). DNA sequences were determined by the Sanger dideoxy method[22] with the same primers used for PCR amplication (above). Five PCR-amplified *cblA* gene clones of PC_{MS}_2323 were generated, three of which were sequenced for confirmatory purpose, with no variation resolved.

PCR Amplification Detection, Using Unique Primer Oligonucleotide Sequences Derived From *cblA* Gene Variants:

Method A:

This method was based upon three 'sets' of VARIANT EXTERNAL (downstream to 3' end of *cblA* gene) anti-sense PCR primers and COMMON INTERNAL sense primers (at 5' end of *cblA* gene). Each set specifically amplifies only one or another of the 3 known variant *cblA* gene sequences, including that uniquely associated with the Toronto/Edinburgh (T/E) epidemic clone.

- To amplify *only* T/E* epidemic lineage clones the following primers were used to generate a 573 base pair (bp) PCR product:

common sense primer:

1

ATGCT GAAAT ACGTT CCGAT

20**

variant anti-sense primer:

573

ATGGT TTTTC AGGA GT

558***

---

* T/E (Toronto/Edinburgh): the sought after, highly transmissible, epidemically spread *cblA*+ clone, for example isolates PC_T-5 clone, PC_T-7, PC_E-1359, PC_E-2315.

** 1-20: in accord with conserved region of *cblA* gene sequence, numbered as in Fig. 4.

*** 558-573: unique sequence beginning 57 bp (base pairs) downstream of 3' end of T/E clone *cblA* gene, i.e. end of *cblA* gene sequence is bp No. 501, thus 501 + 57 = 558.

- To amplify *only* $\underline{MS_p\text{-type}}$* (non-epidemic) the following primers were used to generate a 573 base pair (bp) PCR product:

  common sense primer:
  1
  ATGCT GAAAT ACGTT CCGAT
  20**

  variant anti-sense primer:
  573
  AGGAT TTCCA AAGGA GT
  557***

- To amplify *only* $\underline{MS_w\text{-type}}$**** (non-epidemic) the following primers were used to generate a 573 base pair (bp) PCR product:

  common sense primer:
  1
  ATGCT GAAAT ACGTT CCGAT
  20*

  variant anti-sense primer:
  573
  TGACT TCCGA AGGAC TACT

---

* $\underline{MS_p\text{-type}}$: non-epidemic, negligible transmissibility *cblA*$^+$ strain from Mississippi, for example isolate PC_{MS}-2323.

** In accord with conserved region of *cblA* gene sequence, numbered as Fig. 4.

*** 557-573: unique sequence beginning 56 bp downstream of 3' end of MS_p-type *cblA* gene, i.e. end of *cblA* gene sequence is bp No. 501, thus 501 + 56 = 557.

**** $\underline{MS_w\text{-type}}$: another non-epidemic, negligible transmissibility *cblA*$^+$ strain from Mississippi.

555*

Method B:

This method was based upon three 'sets' of VARIANT EXTERNAL (downstream to 3' end of *cblA* gene) anti-sense PCR primers and *VARIANT* INTERNAL sense primers (at 5' end of *cblA* gene). Each set specifically amplifies only one or another of the 3 known variant *cblA* gene sequences, including that uniquely associated with the Toronto/Edinburgh (T/E) epidemic clone.

- To amplify *only* T/E epidemic lineage clones the following primers were used to generate a 331 base pair (bp) PCR product:

  variant sense primer:

  243

  GACTG CCCCG GCTTT GAA

  260**

  variant anti-sense primer:

  573

  ATGGT TTTTC AGGAG T

  558***

- To amplify only MS_P-type (non-epidemic) the following primers were used to generate a 331 base pair PCR product:

---

* 555-573: unique sequence beginning 54 bp downstream of 3' end of MS_W-type *cblA* gene, i.e. end of *cblA* gene sequence is bp No. 501, thus 501 + 54 = 555.

** 243-260: in accord with *cblA* gene sequence numbering in Fig. 4 for epidemic isolates PC_T-5, PC_T-7, PC_E-1359, PC_E-2315.

*** 558-573: unique sequence beginning 57 bp (base pairs) downstream of 3' end of T/E clone *cblA* gene, i.e. end of *cblA* gene sequence is bp No. 501, thus 501 + 57 = 558.

variant sense primer:
243
GGCCG AGCCG GCGCT GAA
260*

variant anti-sense primer:
573
AGGAT TTCCA AAGGA GT
557**

- To amplify only MS_W-type (non-epidemic) the following primers were used to generate a 331 base pair PCR product:

variant sense primer:
243
GACCG CTCCG TCGCT CAA
260***

variant anti-sense primer:
573
TGACT TCCGA AGGAC TACT
555****

---

* - 243-260: in accord with *cblA* gene sequence numbering in Fig. 4 for non-epidemic Mississippi isolate PC_{MS}-2323.

** 557-573: unique sequence beginning 56 bp downstream of 3' end of MS_P-type *cblA* gene, i.e., end of *cblA* gene sequence is bp No. 501, thus 501 + 56 = 557.

*** 243-260: this (MS_W-type) is a recently discovered variant *cblA* gene sequence carried by another non-epidemic strain of negligible transmissibility from Mississippi.

**** 555-573: unique sequence beginning 54 bp downstream of 3' end of MS_W-type *cblA* gene, i.e. end of *cblA* gene sequence is bp No. 501, thus 501 + 54 = 555.

Method C:

This method was based upon three 'sets' of Internal (within *cblA* gene) PCR primers. Each set specifically amplifies only one or another of the 3 known variant *cblA* gene sequences, including that uniquely associated with the T/E epidemic clone.

METHOD C1: was based on 'common' sense primer and *variant* antisense primers:

- To amplify *only* T/E epidemic lineage clones the following primers were used to generate a 427 base pair (bp) PCR product:

  common sense primer:

  1

  ATGCT GAAAT ACGTT CCGAT

  20*

  variant anti-sense primer:

  427

  CTTCG ACCTT CTTCT GACC

  409**

---

* 1-20: in accord with conserved region of *cblA* gene sequence, numbering as in Fig. 4.

** 427-409: in accord with *cblA* gene sequence numbering in Fig. 4 uniquely for clonal epidemic isolates PC_T-5, PC_T-7, PC_E-1359, PC_E-2315.

- To amplify only MS_P-type (non-epidemic) the following primers were used to generate a 427 base pair (bp) PCR product:

common sense primer:

1

ATGCT GAAAT ACGTT CCGAT

20**

variant anti-sense primer:

427

CAGCG ACAGT TTTCT GGCC

409*

- To amplify only MS_W-type (non-epidemic) the following primers were used to generate a 427 base pair (bp) PCR product:

common sense primer:

1

ATGCT GAAAT ACGTT CCGAT

20**

variant anti-sense primer:

427

CAGCG ACAGT TTTCT GGCC

409***

---

* 427-409: in accord with *cblA* gene sequence numbering in Fig. 4 uniquely for non-epidemic Mississipi isolate PC_{MS}-2323.

** 1-20: in accord with conserved region of *cblA* gene sequence, numbered as in Fig. 4.

*** 427-409: this (MS_W-type) is a recently discovered variant *cblA* gene sequence carried by a non-epidemic strain of negligible transmissibility from Mississippi.

METHOD C2: was based on *variant* sense primers and *variant* antisense primers:

- To amplify *only* T/E epidemic lineage clones the following primers were used to generate a 185 base pair (bp) PCR product:

  variant sense primer:
  243
  GACTG CCCCG GCTTT GAA
  260*

  variant anti-sense primer:
  427
  CTTCG ACCTT CTTCT GACC
  409**

- To amplify *only* MS_P-type (non-epidemic) the following primers were used to generate a 185 base pair (bp) PCR product:

  variant sense primer:
  243
  GGCCG AGCCG GCGCT GAA
  260***

  variant anti-sense primer:
  427
  CTCCG GCCGT CTTCT GTTC

---

* 243-260: in accord with *cblA* gene sequence numbering in Fig. 4 for epidemic isolates PC_T-5, PC_T-7, PC_E-1359, PC_E-2315.

** 427-409: in accord with *cblA* gene sequence numbering in Fig. 4 for epidemic isolates PC_T-5, PC_T-7, PC_E-1359, PC_E-2315.

*** 243-260: in accord with *cblA* gene sequence numbering in Fig. 4 for non-epidemic Mississippi isolate PC_{MS}-2323.

409*

- To amplify *only* MS_W-type (non-epidemic) the following primers were used to generate a 185 base pair (bp) PCR product:

variant sense primer:

243

GACCG CTCCG TCGCT CAA

260**

variant anti-sense primer:

427

CAGCG ACAGT TTTCT GGCC

409***

---

* 427-409: in accord with *cblA* gene sequence numbering in Fig. 4 for non-epidemic Mississippi isolate PC_{MS}-2323.

** 243-260: this (MS_W-type) is a recently discovered variant *cblA* gene sequence carried by another non-epidemic strain of negligible transmissibility from Mississippi.

*** 427-409: this (MS_W-type) is a recently discovered variant *cblA* gene sequence carried by another non-epidemic strain of negligible transmissibility from Mississippi.

The probability of the random occurrence of 16- to 19-base long sequences, such as the specific primers employed in any of the above methods, is once every 4^{16} (or ca. one in 10^9) to 4^{19} (or ca. one in 10^{11}) bp x 2. Given that the size of the *P. cepacia* genome is ca. 7×10^6 bp, the chance occurrence would be remote.

In order to determine whether a given sample contains a transmissible strain of *P. cepcia* any one of the above pairs of primers corresponding to the T/E epidemic lineage is used in standard PCR reactions with DNA from the bacterial strain to be tested.

Following PCR amplification, the reaction products are analyzed for the presence of the appropriate size PCR product using standard methods. For example, in Method A (when the pair of primers for the T/E epidemic lineage is used), the detection of a 573 bp PCR product by electrophoresis on an agarose gel would indicate that the sample contains a strain of the highly transmissible T/E. Whereas, the absence of such a PCR product would indicate otherwise. Similarly in Methods B, C1 and C2, the presence of a strain of the highly transmissible T/E would be confirmed by PCR products of 331 bp, 427 bp and 185 bp, respectively.

Hybridization-specific detection, based on 'probes' derived from unique regions of variant *cblA* gene sequences:

Each unique probe sequence shown below can be generated by restriction endonuclease double digestion of one or another of the three *cblA* gene variants, based on:

(i) common *Sma*I site at bp (base pair) 273-278 (C CCGGG), flanking 3' end of *cblA* probe

(ii) common *Bsr*I site at bp 371-376 (AAC TGG), flanking 5' end of *cblA* probe

the *cblA*-specific DNA probes:

- Hybridization with *only* T/E epidemic lineage clones

Unique T/E *cblA* gene 'probe' sequence:

273
C CCGGG CGCGG CGGAA ATTCC GCTGT CGGTC AAGCT
TGGCG AAACC GAGCT GACCA CCACG GCCGC GACGC
TGAAG ACCGC AGAGC TCTTC CCCGG CGAAC TGG
376*

---

* 273-376: in accord with *cblA* gene sequence numbering in Fig. 4 for epidemic isolates PC_T-5, PC_T-7, PC_E-1359, PC_E-2315.

- Hybridization with *only* MS_P-type (non-epidemic) strains

  Unique MSP-type *cblA* gene 'probe' sequence:

  273

  C CCGGG CGCCA AGGAG ATTCC GCTGG CGGTC AAGCT
  GGGCA CCACC GCGCT GAGCA CCACG GCGAC GACGC
  TGAAG GCGTC GGAGA TCTTT ACCGG CGAAC TGG

  376*

- Hybridization with only MSW-type (non-epidemic) strains

  Unique MS_W-type *cblA* gene 'probe' sequence:

  273

  C CCGGG TGCGG CCGAA ATTCC GTTGG CTGTC AAACT
  GGGCG ACACC GAGTT GAGCA CGACG TCGGC GACCC
  TGAAG GCCGC GGAAC TCTTT CCCGG CGAAC TGG

  376**

The probes for the T/E lineage disclosed above can be used in standard hybridization methods, such as colony hybridization assays or Southern blot transfer, to detect the presence of the highly transmissible *P. cepacia* strain of the T/E lineage.

---

* 273-376: in accord with *cblA* gene sequence numbering in Fig. 4 for non-epidemic Mississippi isolate PC_{MS}-2323.

** 273-376: this (MS_W-type) is a recently discovered variant *cblA* gene sequence carried by another non-epidemic strain of negligible transmissibility from Mississippi.

RESULTS

The above-described methods were used to characterize the epidemiological-relatedness of the 15 Toronto isolates expressing mucin-binding Cbl pili. To investigate the genetic relationship between these isolates and those found elsewhere, we included clinical and environmental isolates as well as 78 strains from the seven other CF centers cited above that were *cblA*$^-$. At the time the report of RFLP-identical *P. cepacia* isolates transmitted among patients at an Edinburgh CF center appeared and we obtained the involved strains[3] to include in this phylogenetic characterization. See **Fig. 1A and 1B.** Profiles in **lanes 9-17** of both figures depict typical polymorphic patterns resolved for isolates from different CF centers. For these isolates mean D (Dice coefficient of similarity)[12] for any pair by PFGE-resolved chromosomal macro-RFLP profile was 0.14 ± 0.07 **(Fig. 1A)**, a level of diversity not significantly different from that found previously among eight independently isolated ATCC environmental and clinical control isolates[9]. A similar degree of chromosomal RFLP variability was found between the other CF-associated isolates from the seven CF centers (results not shown), confirming that these are epidemiologically distinct strains with RFLP variability not significantly different from that of the random collection of ATCC strains ($0.1>p>0.05$). Despite the lower discriminatory power of ribotyping[9], a similar degree of phylogenic relationship among these CF-associated isolates is apparent in **Fig. 1B.**

The heterogeneity of the RFLP profiles of the isolates from the seven CF centers (e.g. **lanes 9-17, Fig. 1A and 1B**) is similar to that described in our previous study involving multiple isolates from 23 patients at the UNC CF center[9]. This degree of variability contrasts markedly with the two closely-related, conserved RFLP patterns found for the 15 *cblA*⁻ encoding Toronto CF center isolates (**lanes 1-4, Fig. 1A and 1B**). Here, by examining both PFGE and ribotype RFLP profiles, the coefficient of similarity amongst the Toronto isolates proved to be very high, with PFGE **D** = 0.95±0.03 and *rrn* **D** = 0.87±0.09. This contrasted with (a) the mean **D** values amongst isolates from the other seven CF centers which was very low (e.g. **lanes 9-17, Fig. 1A and 1B**), and (b) the mean **D** between the Toronto isolates and the other CF center isolates which was also very low: PFGE **D** = 0.20±0.07, *rrn* **D** = 0.39±0.09. These findings strongly suggest that all 15 of the Toronto CF center isolates were members of a unique lineage associated with an epidemic.

Displayed in **lanes 5-8** of **Fig. 1A and 1B** are *P. cepacia* PFGE and ribotype RFLP profiles of isolates from CF patients at the Edinburgh CF center[3]. The RFLP profiles displayed in **Fig. 1A and 1B** support this picture, as **D** for any analyzed pair by either type of RFLP profile was very high (PFGE **D** = 0.98±0.02, *rrn* **D** = 1.0). Further, pair-wise comparison of the Edinburgh strains to the closely-related Toronto CF center strains (**lanes 1-4** of **Fig. 1A and 1B**) likewise produced robust **D** values (PFGE **D** = 0.97±0.03, *rrn* **D** = 0.90±0.04), strongly suggesting that the Edinburgh and Toronto isolates are members of the same unique lineage, despite the Atlantic Ocean barrier.

Based on ribosomal RNA (*rrn*) RFLP profiles, phylogenic relationships of the 133 isolates described above and in the Methods section were determined with the neighbor-joining method[13]. Confidence intervals on the tree topology were estimated by bootstrapping analysis[14]. See **Fig. 2.** The resultant phylogenic tree indicates that: (i) the cluster of Toronto and Edinburgh isolates comprise a single, clonally-related lineage, only distantly related to all other isolates, and (ii) the remaining, independently isolated strains from other CF centers are as distantly-related to one another as they are either to the Toronto/Edinburgh clusters or to the independently isolated non-CF clinical and environmental strains.

We then examined the epidemic Edinburgh isolates to see if, like the Toronto strains[4], they expressed Cbl pili and encoded the *cblA* gene. Phenotypic survey using electron microscopy revealed that these highly transmissible strains expressed appendage pili that were structurally equivalent to those expressed by all of the *cblA*$^+$ Toronto isolates[4, 11]. See **Fig. 3.** Genotypic survey was carried out by stripping *rrn*-probe from an *Eco*RI chromosomal digest membrane (**Fig. 1B**) followed by hybridization with a *cblA* probe. **Fig. 1C (lanes 1-8)** indicates that four highly transmissible Edinburgh isolates as well as the closely-related Toronto clones encode *cblA*.

Because the implications of our studies have the potential to directly affect clinical management of some 70,000 CF patients in Europe and North America, we used DNA sequence analysis to test the RFLP-based conclusions that isolates from Toronto and Edinburgh were clonal.

Primers were synthesized from the *cblA* sequence of the Toronto isolate PC_{T}-74 and used for PCR-based amplification of the *cblA* gene from isolates to be characterized. Resultant PCR products were then cloned and sequenced (see Methods). Complete *cblA* sequences were thus obtained from isolates with the two slightly variant though closely-related RFLP profiles typical of the 15 Toronto CF center isolates (**Fig. 1A and 1B, lanes 1-4**), the two slightly variant though closely-related RFLP profiles typical of the 13 Edinburgh CF center isolates (**Fig. 1A and 1B, lanes 5-8**), and the significantly variant Jackson Mississippi CF-associated strain PC_{MS}-2323 (**Fig. 1A and 1B, lane 17**). Comparison of these five sequences indicated that the chromosomally-encoded, 501 bp *cblA* pilin subunit structural gene carried by the closely-related isolates from the Toronto and Edinburgh CF centers was invariant in sequence. In contrast, the *cblA* gene encoded by the distantly-related Jackson Mississippi strain PC_{MS}-2323 exhibited polymorphism at the sequence level, with changes in 60 bp of the 501 bp sequence (88% identity; see **Figure 4**). The perfect conservation of the *cblA* pilin sequence among isolates from multiple patients over four years at the Toronto and Edinburgh centers is precisely what would be expected for epidemic transmission of a highly infectious clone. Likewise, variability of *cblA* encoded by the Mississippi CF center isolate is in accord with that expected for a distantly-related isolate.

CONCLUSIONS

Consideration of these findings leads us to conclude: (i) isolates of *P. cepacia* are not equally

transmissible between CF patients; rather, there exists at least one significantly divergent, highly transmissible clonal lineage plus numerous moderately heterogeneous lineages of negligible transmissibility (see **Fig. 2**), and (ii) the highly transmissible lineage identified is responsible for epidemics at North American and British CF centers. This was most likely due to an as yet unidentified transatlantic transmission, possibly summer camp attendance.

Based solely on *rrn* RFLP profiles or anecdotal evidence, additional reports exist suggesting the occurrence of *P. cepacia* transmission at CF centers in Philadelphia[2] and Cleveland[17], respectively. We characterized strains involved with both of these putative outbreaks (see Methods section) and found that by neither ribotype or macro-chromosomal RFLPs profile did the prototypic RFLP fingerprints of the putatively epidemic strains from either center appear similar to one another (mean $D \leq 0.3$), nor to the unique, highly transmissible lineage involved with the Toronto and Edinburgh CF patients. Nonetheless, highly conserved RFLP profiles (mean $D \geq 0.85$) within the individual outbreaks did support a picture of epidemic transmission within each of the two centers. When these isolates were further characterized, hybridization-based survey for the presence of the *cblA* pilin gene proved negative for all 35 of the involved strains. These results suggest that there may exist *P. cepacia* lineages of high transmissibility other than the $cblA^+$ clone that we have identified.

While we have hereinbefore described a number of embodiments of this invention, it is apparent that our basic disclosures can be altered to provide other

embodiments which utilize the methods and compositions of this invention.

To the extent the following references include protocols or materials employed in the methods discussed herein, they are incorporated herein by reference.

1. LiPuma, J., Mortensen, J., Dasen, S. & Stull, T. Ribotype analysis of *P. cepacia* from cystic fibrosis treatment centers. J. Pediatrics. 113, 859-862 (1988).

2. LiPuma, J., Dasen, S., Nielson, D., Stern, R. & Stull, T. Person-to-person transmission of *P. cepacia* between patients with cystic fibrosis. Lancet. 336, 1094-1096 (1990).

3. Govan, J. R. W. et al. Evidence for transmission of *Pseudomonas cepacia* by social contact in cystic fibrosis. Lancet. 342, 15-19 (1993).

4. Sajjan, S. et al. Cable (Cbl) type II pili of cystic fibrosis-associated *Burkholderia (Pseudomonas) cepacia*: nucleotide sequence of the *cblA* major subunit pilin gene and novel morphology of the assembled appendage fibers. J. Bacter. 177, 1030-1038 (1995).

5. Tablan, O. et al. *Pseudomonas cepacia* colonization in patients with cystic fibrosis: Risk factors and clinical outcome. J. Pediatr. 107, 382-387 (1985).

6. Thomassen, M. J., Demko, C., Klinger, J. & Stern, R. *Pseudomonas cepacia* colonization among patients with cystic fibrosis - A new opportunist. Am. Rev. Respir. Dis. 131, 791-796 (1985).

7. Ramirez, J. C.,et al. Bilateral lung transplantation for cystic fibrosis. J. Thor. Cardiovasc. Surg. 103, 287-294 (1992).

8. Snell, G. I., deHoyos, A., Krajden, M., Winton, T. & Maurer, J.R. *Pseudomonas cepacia* in lung transplant recipients with cystic fibrosis. Chest. 103, 466-471 (1993).

9. Steinbach, S. et al. Transmissibility of *Pseudomonas cepacia* infection in clinic patients and lung-transplant recipients with cystic fibrosis. The New England J. of Med. 331, 981-987 (1994).

10. Anderson, D. J., Kuhns, J.S., Vasil, M.L., Gerding, D.N. & Janoff, E.N. DNA fingerprinting by pulsed field gel electrophoresis and ribotyping to distinguish *Pseudomonas cepacia* isolates from a nosocomial outbreak. J. Clin. Microbiol. 29, 648-649 (1991).

11. Goldstein, R. et al. Structurally variant classes of pilus appendage fibers coexpressed from Burkholderia (Pseudomonas) cepacia. J. Bacteriol. 177, 1039-1052 (1995).

12. Dice, L. Measures of the amount of ecological association between species. Ecology. 26, 297-302 (1945).

13. Saitou, N. & Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425 (1987).

14. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 39, 783-791 (1985).

15. Nelson, K. & Selander, R. Evolutionary genetics of the proline permease (putP) and the control region

of the proline utilization operon in populations of Salmonella and E. coli. J. Bacteriol. 174, 6886-6895 (1992).

16. Smith, N. & Selander, R. Sequence invariance of the antigen-coding central region of the phase 1 flagellar filament gene (fliC) among strains of Salmonella typhimurium. J. Bacteriol. 172, 603-609 (1990).

17. Thomassen, M. J., C. Demko, C. Doershuk, R. C. Stern & J. Klinger. *Pseudomonas cepacia*: Decrease in colonization in patients with cystic fibrosis. Am. Rev. Respir. Dis. 134, 669-671 (1986).

18. Sajjan, U., Corey, M., Karmali, M. & Forstner, J. Binding of *Pseudomonas cepacia* to normal human intestinal mucin and respiratory mucin from patients with cystic fibrosis. J. Clin. Invest. 89, 648-556 (1992).

19. Sajjan, S. & Forstner, J. Identification of the mucin-binding adhesin of *P. cepacia* isolated from patients with cystic fibrosis. Infect. Immun. 60, 1434-1440 (1992).

20. Vinograd, J. & Hearst, J.E. Equilibrium sedimentation of macromolecules and viruses in a density gradient. Fortschritte der Chemie Organischer Naturstoffe. XX, 372-422 (1962).

21. Mullis, K. & Faloona, F. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Meth. Enzymol. 155, 335-350 (1987).

22. Sanger, F., Nicklen, S. & Coulson, A. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. (USA). 74, 5463-5467 (1977).

23. Prevost, G., Jaulhac, B. & Piemont, Y. DNA fingerprinting by PFGE is more effective than ribotyping in distinguishing among MRSA isolates. J. Clin. Microbiol. 30, 967-973 (1992).

24. Rabkin, C., Jarvis, W., Anderson, R., Stull, T. & Woods, D. *P. cepacia* typing systems: Collaborative study to assess their potential in epidemiologic investigations. Rev. Infect. Dis. 11, 600-607 (1989).

25. Nei, M. & Li, W. H. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA. 76, 5269-5273 (1979).

26. Arthur, M. et al. Restriction fragment length polymorphisms among uropathogenic Escherichia coli: pap-related sequences compared with *rrn* operons. Infect. Immun. 58, 471-479 (1990).

27. Arthur, M. et al. Molecular epidemiology of adhesin and hemolysin virulence factors among uropathogenic Escherichia coli. Infect. Immun. 57, 303-313 (1989).

28. Berg, K., Squires, C. & Squires, C. *In vivo* translation of a region within the *rrnB* 16S rRNA gene of *Escherichia coli*. J. Bacteriol. 169, 1691-1701 (1987).

29. Birnboim, H. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol. 100, 243-255 (1983)

30. Feinberg, A. & Vogelstein, B. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137, 266-267.

We Claim:

1. A method for detecting in a sample the presence of a strain of a transmissible lineage of *Pseudomonas cepacia* comprising the step of:

analyzing the sample for restriction fragment length polymorphisms (RFLPs) linked to a strain known to be of a transmissible lineage of *Pseudomonas cepacia*.

2. The method according to claim 1 wherein the sample is analyzed by ribotyping.

3. The method according to claim 1 wherein the sample is analyzed by DNA fingerprinting using pulsed field gel electrophoresis.

4. The method according to any one of claims 1-3 wherein the strain known to be of a transmissible lineage of *Pseudomonas cepacia* expresses cable adhesin type II pili.

5. The method according to claim 4 wherein the strain of *Pseudomonas cepacia* is of the highly transmissible Toronto/Edinburgh lineage.

6. The method according to claim 5 wherein the strain is selected from the group consisting of: PC_T_5, PC_T_7, PC_T_19, PC_T_25, PC_E_509, PC_E_1359, PC_E_1392 and PC_E_2315.

7. A method for detecting in a sample the presence of a strain of a transmissible lineage of *Pseudomonas cepacia* by polymerase chain reaction (PCR)

using one or more pairs of oligonucleotide primers, said primers having nucleotide sequences identical to portions of a gene encoding a 17 kDa major subunit pilin protein of the cable adhesin type II *P. cepacia* pili.

8. A method for detecting in a sample the presence of a strain of a transmissible lineage of *Pseudomonas cepacia* by polymerase chain reaction (PCR) using one or more pairs of oligonucleotide primers selected from the group consisting of:

(a) common sense primer:

1
ATGCT GAAAT ACGTT CCGAT
20

variant anti-sense primer:

573
ATGGT TTTTC AGGA GT
558;

(b) *variant* sense primer:

243
GACTG CCCCG GCTTT GAA
260

variant anti-sense primer:

573
ATGGT TTTTC AGGAG T
558;

(c) common sense primer:
1
ATGCT GAAAT ACGTT CCGAT
20
variant anti-sense primer:
427
CTTCG ACCTT CTTCT GACC
409; and

(d) variant sense primer:
243
GACTG CCCCG GCTTT GAA
260
variant anti-sense primer:
427
CTTCG ACCTT CTTCT GACC
409.

9. A method for detecting in a sample the presence of a strain of a transmissible lineage of *Pseudomonas cepacia* by DNA hybridization using a DNA probe having the DNA sequence:

273
C CCGGG CGCGG CGGAA ATTCC GCTGT CGGTC AAGCT
TGGCG AAACC GAGCT GACCA CCACG GCCGC GACGC
TGAAG ACCGC AGAGC TCTTC CCCGG CGAAC TGG
376.

10. A pair of oligonucleotide primers selected from the group consisting of:

(a) common sense primer:
1
ATGCT GAAAT ACGTT CCGAT
20
variant anti-sense primer:
573
ATGGT TTTTC AGGA GT
558;

(b) *variant* sense primer:
243
GACTG CCCCG GCTTT GAA
260
variant anti-sense primer:
573
ATGGT TTTTC AGGAG T
558;

(c) common sense primer:
1
ATGCT GAAAT ACGTT CCGAT
20
variant anti-sense primer:
427
CTTCG ACCTT CTTCT GACC
409; and

(d) variant sense primer:
243
GACTG CCCCG GCTTT GAA
260
variant anti-sense primer:
427
CTTCG ACCTT CTTCT GACC
409.

11. A pair of oligonucleotide primers selected from the group consisting of:

(a) common sense primer:
1
ATGCT GAAAT ACGTT CCGAT
20
variant anti-sense primer:
573
AGGAT TTCCA AAGGA GT
557;

(b) common sense primer:
1
ATGCT GAAAT ACGTT CCGAT
20
variant anti-sense primer:
573
TGACT TCCGA AGGAC TACT
555;

(c) *variant* sense primer:
243
GGCCG AGCCG GCGCT GAA
260

variant anti-sense primer:
573
AGGAT TTCCA AAGGA GT
557;

(d) *variant* sense primer:
243
GACCG CTCCG TCGCT CAA
260
variant anti-sense primer:
573
TGACT TCCGA AGGAC TACT
555;

(e) common sense primer:
1
ATGCT GAAAT ACGTT CCGAT
20
variant anti-sense primer:
427
CTCCG GCCGT CTTCT GTTC
409;

(f) common sense primer:
1
ATGCT GAAAT ACGTT CCGAT
20
variant anti-sense primer:
427
CAGCG ACAGT TTTCT GGCC
409;

(g) *variant* sense primer:
243
GGCCG AGCCG GCGCT GAA
260
variant anti-sense primer:
427
CTCCG GCCGT CTTCT GTTC
409; and

(h) *variant* sense primer:
243
GACCG CTCCG TCGCT CAA
260
variant anti-sense primer:
427
CAGCG ACAGT TTTCT GGCC
409.

12. A DNA probe having the DNA sequence:
273
C CCGGG CGCGG CGGAA ATTCC GCTGT CGGTC AAGCT
TGGCG AAACC GAGCT GACCA CCACG GCCGC GACGC
TGAAG ACCGC AGAGC TCTTC CCCGG CGAAC TGG
376.

13. A DNA probe having a DNA sequence selected from the group consisting of:

(a) C CCGGG CGCCA AGGAG ATTCC GCTGG CGGTC AAGCT GGGCA CCACC GCGCT GAGCA CCACG GCGAC GACGC TGAAG GCGTC GGAGA TCTTT ACCGG CGAAC TGG; and

(b) C CCGGG TGCGG CCGAA ATTCC GTTGG CTGTC AAACT GGGCG ACACC GAGTT GAGCA CGACG TCGGC GACCC TGAAG GCCGC GGAAC TCTTT CCCGG CGAAC TGG.

14. A diagnostic kit comprising at least one pair of oligonucleotide primers of claim 10.

15. A diagnostic kit comprising the DNA probe of claim 12.

16. An isolated DNA molecule encoding the 17 kDa major subunit pilin protein of the cable adhesin type II *P. cepacia* pilus.

17. The DNA molecule according to claim 15 having a DNA sequence selected from the two DNA sequences depicted in Figure 4.

18. A recombinant DNA molecule comprising a DNA molecule of claim 16.

19. A unicellular host transformed with a recombinant DNA molecule according to claim 18.

20. A method of producing a polypeptide comprising the step of culturing a unicellular host transformed with a recombinant DNA molecule according to claim 18.

21. A polypeptide produced according to the method of claim 20.

22. An antibody directed against a polypeptide of claim 21.

23. A diagnostic kit comprising a polypeptide of claim 21.

24. A diagnostic kit comprising an antibody of claim 22.

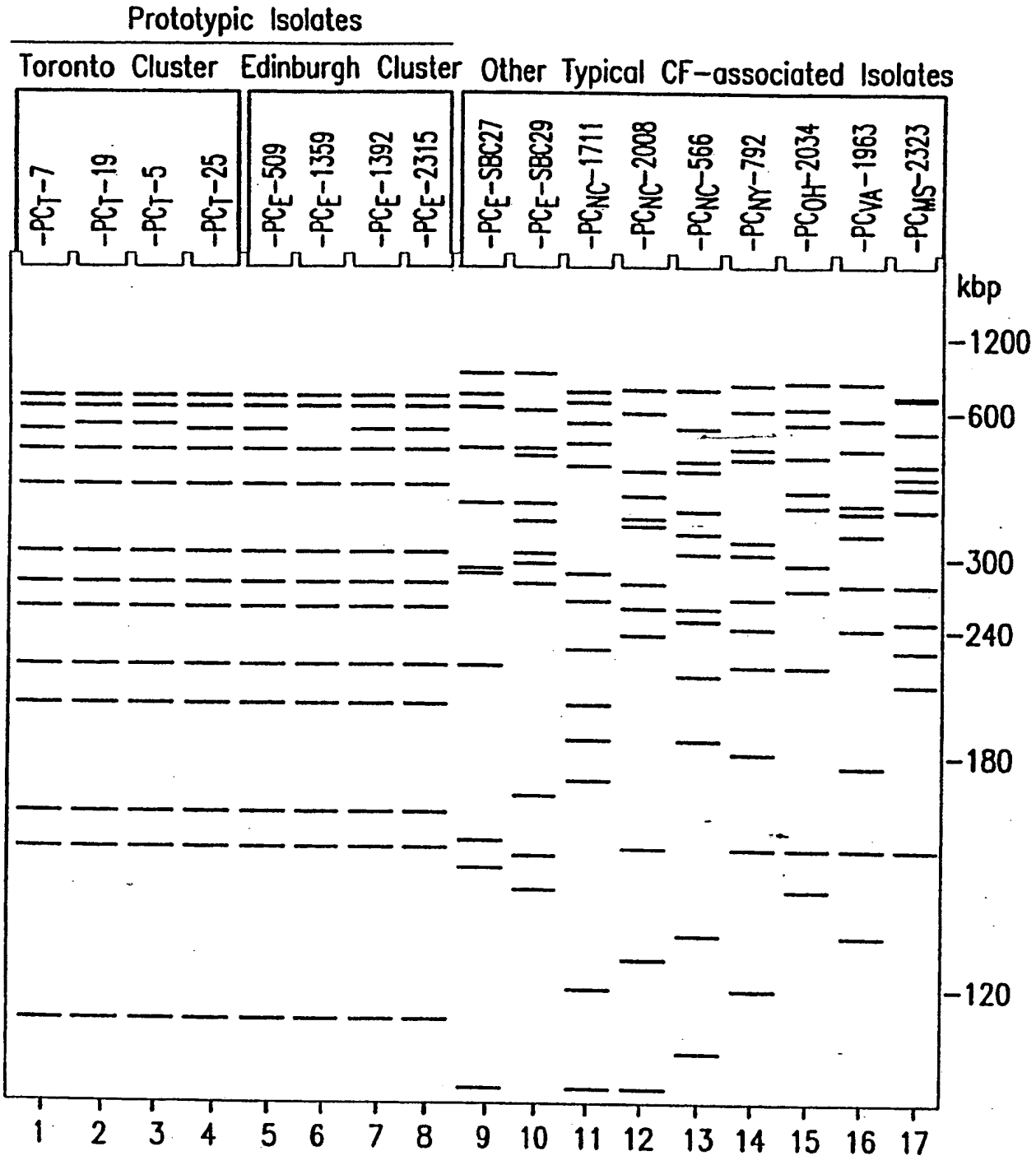
Prototypic Isolates
Toronto Cluster
Edinburgh Cluster
Other Typical CF-associated Isolates
-PCT-7
-PCT-19
-PCT-5
-PCT-25
-PCE-509
-PCE-1359
-PCE-1392
-PCE-2315
-PCE-SBC27
-PCE-SBC29
-PCNC-1711
-PCNC-2008
-PCNC-566
-PCNY-792
-PCOH-2034
-PCVA-1963
-PCMS-2323
kbp
1200
600
300
240
180
120
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

FIG.1A

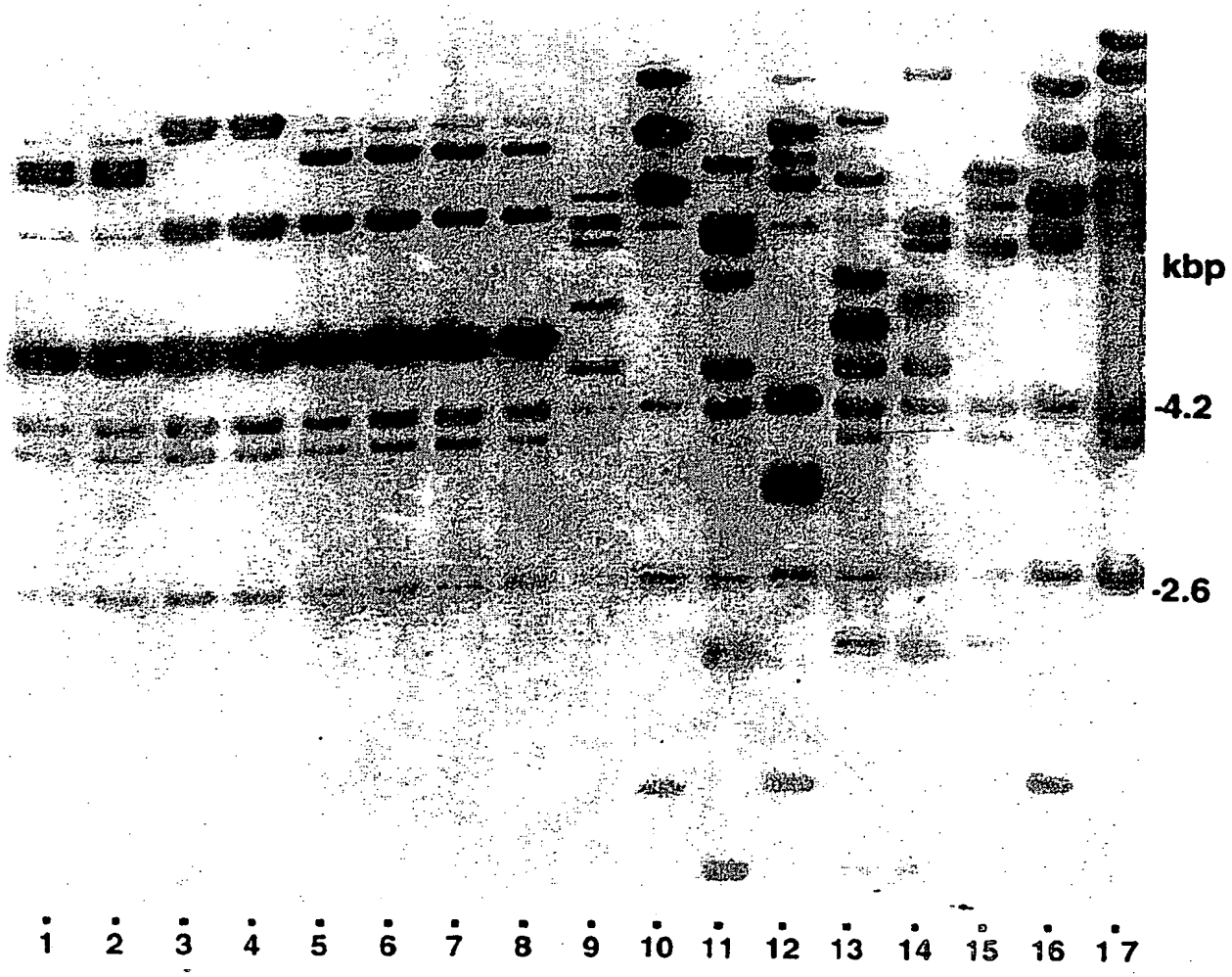
kbp
-4.2
-2.6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

FIG.1B

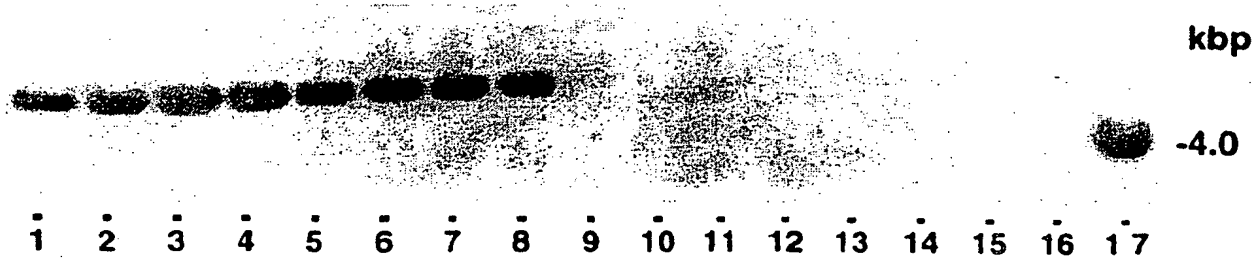
kbp
-4.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

FIG.1C

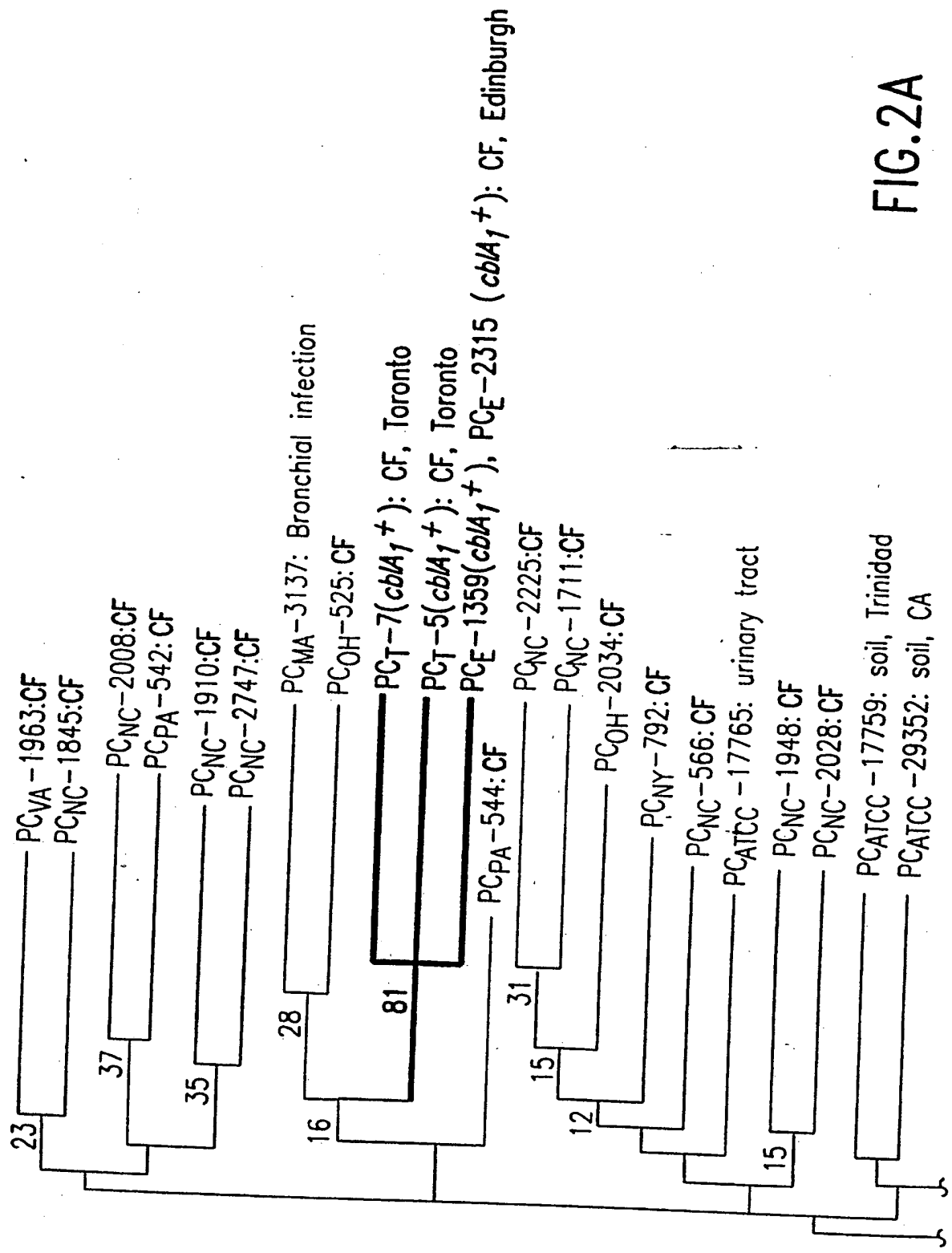
PCVA-1963:CF
PCNC-1845:CF
PCNC-2008:CF
PCPA-542: CF
PCNC-1910:CF
PCNC-2747:CF
PCMA-3137: Bronchial infection
PCOH-525: CF
PCT-7(cblA1+): CF, Toronto
PCT-5(cblA1+): CF, Toronto
PCE-1359(cblA1+), PCE-2315 (cblA1+): CF, Edinburgh
PCPA-544: CF
PCNC-2225:CF
PCNC-1711:CF
PCOH-2034:CF
PCNY-792: CF
PCNC-566: CF
PCATCC-17765: urinary tract
PCNC-1948: CF
PCNC-2028: CF
PCATCC-17759: soil, Trinidad
PCATCC-29352: soil, CA
23
37
35
28
81
31
15
12
15

FIG.2A

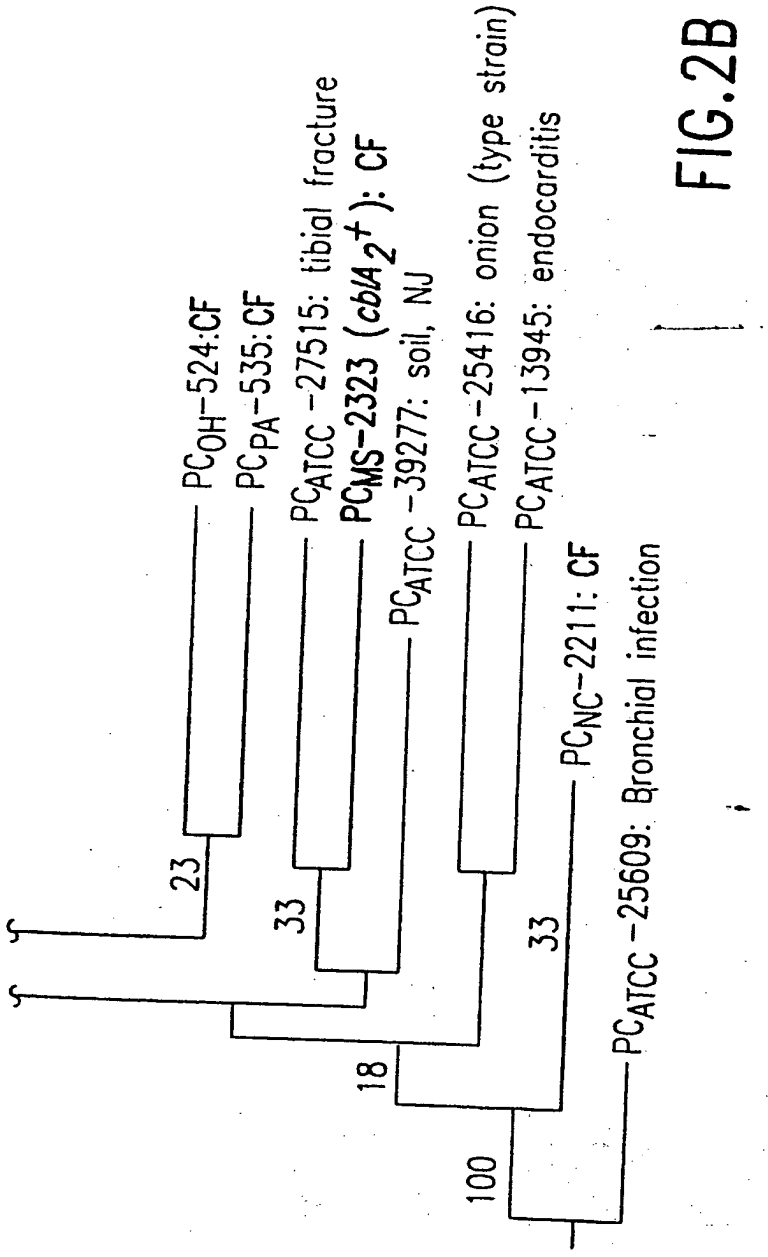

FIG.2B

FIG.3

```
  1 ATGCTGAAATACGTTCCGATCGCTGCTGCTGCTCTGATGTCGATGTCGGC 50
    ||||||||||||||||||||||  ||| || || ||||||||||||||||
  1 ATGCTGAAATACGTTCCGATCCGTGCCGCCGCCCTGATGTCGATGTCGGC 50

 51 TTACGCCGTCCAGAAGGACATTACCGTCACCGCCAACGTCGACACGACGC 100
    | ||||||| |||||||||| |||||||||||||||||||||||||||||
 51 TCACGCCGTGCAGAAGGACATCACCGTCACCGCCAACGTCGACACGACGC 100

101 TCGAAATGCTGTCGGCGGACGGCTCGGCACTGCCGACGACCATGCAGATG 150
    ||||  |||||||||||||||||  |||||  ||||||||||||||||||
101 TCGAGATGCTGTCGGCGGACGGTTCGGCGCTGCCGACGACCATGCAGATG 150

151 CAATATCTGCCGGGTACGGGTCTTCAGGCAGCTGTAGTGAACACGAAGAT 200
    ||  ||||||||||||  |||||||  |||||  ||||  ||||||||||||||||
151 CAGTATCTGCCGGGCTCGGGTCTCCAGGCGGCTGAAGTGAACACGAAGAT 200

201 CTTCACGAACGACAAGGCAAAGGATCTGCAGATCCGCCTCGCGACTGCCC 250
    ||||||||||||||||||  ||||||||||||||||||||||||| | |  |
201 CTTCACGAACGACAAGGCGAAGGATCTGCAGATCCGCCTCGCGGCCGAGC 250

251 CGGCTTTGAAGAACCAGACGAGCCCGGGCGCGGCGGAAATTCCGCTGTCG 300
    ||||  |||||||||||||||||||||||||  ||| |||||||||  ||
251 CGGCGCTGAAGAACCAGACGAGCCCGGGCGCCAAGGAGATTCCGCTGGCG 300

301 GTCAAGCTTGGCGAAACCGAGCTGACCACCACGGCCGCGACGCTGAAGAC 350
    ||||||||  |||   ||||  |||||  ||||||||   |||||||||| |
301 GTCAAGCTGGGCACCACCGCGCTGAGCACCACGGCGACGACGCTGAAGGC 350

351 CGCAGAGCTCTTCCCCGGCGAACTGGCACAAGGTTCGAACGTGCTGGCGC 400
    | |||  ||||  ||||  |||||||| ||  || |||||||||||  |||
351 GTCGGAGATCTTTACCGGTGAACTGGCGCAGGGCTCGAACGTGCTGCCGC 400

401 TGTCGATCGGTCAGAAGAAGGTCGAAGCCGTCACGGCGTCCGGCAGCTAC 450
    |||||||||  |||||||  ||  ||  ||||  |||||||  ||||||||||
401 TGTCGATCGAACAGAAGACGGCCGGAGCCATCACGGCTGCCGGCAGCTAT 450

451 CAGGGCCTCGTCAGCGTGATCGTCACGCAGAGCGCGGCTTCGGGTAGCTA 500
    ||  ||||||||||||||||||||||||||||||||||  ||||||||||
451 CAAGGCCTCGTCAGCGTGATCGTCACGCAGAGCGCGGCATCGGGTAGCTA 500

501 A 501
    |
501 A 501
```

FIG. 4

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, C12N 15/31, C07K 14/21, 16/12 — A2

(11) International Publication Number: WO 97/01647

(43) International Publication Date: 16 January 1997 (16.01.97)

(21) International Application Number: PCT/US96/11132

(22) International Filing Date: 28 June 1996 (28.06.96)

(30) Priority Data:
60/000,828 28 June 1995 (28.06.95) US

(71) Applicant: TRUSTEES OF HEALTH & HOSPITALS OF THE CITY OF BOSTON, INC. [US/US]; 6th floor, 1010 Massachusetts Avenue, Boston, MA 02119 (US).

(72) Inventor: GOLDSTEIN, Richard; 12 Howland Street, Cambridge, MA 02138 (US).

(74) Agents: PIERRI, Margaret, A. et al.; Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: DNA SEQUENCES FOR IDENTIFYING HIGHLY TRANSMISSIBLE LINEAGES OF PSEUDOMONAS (BURKHOLDERIA) CEPACIA

(57) Abstract

The present invention provides DNA based fingerprints and DNA sequences for indentifying highly transmissible lineages of *Pseudomonas cepacia*. More specifically, the present invention provides genetic band patterns or DNA based fingerprints for identifying highly transmissible lineages of *Pseudomonas cepacia* produced by: (a) ribotyping analyses, i.e., the determination of restriction fragment length polymorphisms (RFLPs) associated with the multicopy RNA operon (rrn); and (b) pulsed field gel electrophoresis (PFGE)-based resolution of chromosomal macro-RFLP patterns. Also provided are unique primer oligonucleotide sequences and DNA probes derived from variants of a gene, encoding a 17-KDa major subunit pilin protein (cblA) of a cystic fibrosis-associated *Pseudomonas cepacia*. The invention also discloses methods and diagnostic kits for identifying highly transmissible lineages of *Pseudomonas cepacia* that utilize the above DNA based fingerprints, oligonucleotide primers and DNA probes. Another object of this invention is to provide isolated DNA molecules encoding a 17-KDa major subunit pilin protein of certain cystic fibrosis-associated *Pseudomonas cepacia* strains, as well as recombinant DNA molecules, transformed hosts and methods for the production of that protein. Also contemplated are antibodies to the 17-KDa major subunit pilin protein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia
AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam